

Pricing in Social Networks with Negative Externalities *

Zhigang Cao Xujin Chen Xiaodong Hu Changjun Wang

Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100190, China

{zhigangcao, xchen, xdhu, wcj}@amss.ac.cn

Abstract

We study the problems of pricing an indivisible product to consumers who are embedded in a given social network. The goal is to maximize the revenue of the seller. We assume impatient consumers who buy the product as soon as the seller posts a price not greater than their valuations of the product. The product's value for a consumer is determined by two factors: a fixed consumer-specified intrinsic value and a variable externality that is exerted from the consumer's neighbors in a linear way. We study the scenario of negative externalities, which captures many interesting situations, but is much less understood in comparison with its positive externality counterpart. We assume complete information about the network, consumers' intrinsic values, and the negative externalities. The maximum revenue is in general achieved by iterative pricing, which offers impatient consumers a sequence of prices over time.

We prove that it is NP-hard to find an optimal iterative pricing, even for unweighted tree networks with uniform intrinsic values. Complementary to the hardness result, we design a 2-approximation algorithm for finding iterative pricing in general weighted networks with (possibly) nonuniform intrinsic values. We show that, as an approximation to optimal iterative pricing, single pricing works rather well for many interesting cases, such as forests, Erdős-Rényi networks and Barabási-Albert networks, although its worst-case performance can be arbitrarily bad.

Keywords: Pricing, Algorithmic Game Theory, Social Networks, Negative Externalities, Random Networks

1 Introduction

People interact with and influence each other to a degree that is beyond most of us can imagine. The magnitude of this connection has been upgraded to a brandnew level by the proliferation of online SNS (Social Network Services, e.g. Facebook, Twitter, Google Plus, and SinaWeibo). Numerous business opportunities are being incubated by this upgrading. Yet, its consequences are far from being fully unfolded or understood, leaving many fascinating questions for scientists in a variety of disciplines to answer. One incredible fact in the SNS era is that we are now able to know the complete network of who is connected with whom. Network marketing and pricing, with the assistance of *big data*, could be much more precise and flexible than traditional counterparts, and are attracting increasing attention from both industry and academia. In this paper, we study, from an algorithmic point of view, how a monopolist seller should price to the consumers connected by a known social network.

Consumption is never a completely private thing. As opposed to standard economic settings, the utilities that a consumer obtains from consuming many kinds of goods, are not determined merely by his/her private needs and the functions and qualities of the goods, but also greatly affected by the consumptions of his/her social network neighbors. For example, the reason that we wear clothes is not only to cover ourselves

*Supported in part by NNSF of China under Grant No. 11222109, 71101140 and 11471326, 973 Project of China under Grant No. 2011CB80800 and 2010CB731405, and CAS Program for Cross & Cooperative Team of Science & Technology Innovation.

from cold, but usually also to make other people think that we look great and unique. This social side of consumption is becoming more and more prominent with the unification of E-commerce and SNS. It is now very convenient for us to share with our friends our shopping results. By clicking one more button at the time we pay for the skirt online, all our Twitter friends know immediately the complete information of this skirt. This effect could be much stronger and faster than face-to-face sharing. Our ladybros may think the skirt terrific and get one too, or oppositely, they may prefer later a different style to avoid outfit clash. The former case is typical *positive externality*: the incentive that a consumer buys a product increases as more and more of his/her social network neighbors buy the product. The latter opposite scenario, the incentive decreases when more neighbors have the product, is referred to as *negative externality*, which is the focus of this paper. Positive externalities are prevalent in many aspects of the society and have been extensively studied under various academical terms (herd behavior, Matthew effect, strategic complements, and viral marketing, to name a few). Negative externalities, in contrast, although widely exist, are much less studied.

Pricing with negative externalities. We concentrate on the negative externality among consumers of consuming a single kind of product, which is usually luxury or fashionable one. An important reason that a consumer buys this product is to showoff in front of his/her friends (also referred to as *invidious consumption* in literature). Naturally, a consumer buys the product if the price is not higher than his/her (total) value of the product, which is the sum of his/her constant *intrinsic value* and varying *external value*. We propose and study the typical network pricing model, where the external value is the (weighted) number of people to whom the consumer can showoff (i.e. his/her social network neighbors who do not possess this product). We study, to obtain a maximum revenue, how a monopolist seller should price such a product with negative externality to consumers connected by a link-weighted social network, where the revenue is the total payment the seller receives, and the nonnegative integer link weights represent the influences between consumers. While, with the help of SNS, the knowledge of social network structures and real-time externalities is available, consumers' intrinsic values might be known in complete information scenarios, or partially known in incomplete information scenarios. This paper addresses the pricing problems for revenue maximization in complete information scenarios. Our study falls into the framework of *uniform pricing*, where at any time point the same take-it-or-leave price is offered (posted) to all consumers who have not bought the product. The seller adopts a strategy of *iterative pricing* – posting different prices sequentially at discrete time points, to maximize her revenue (we assume that production costs are zero). We also assume that the consumers are *myopic* (a.k.a. *impatient*) in the sense that they, when making purchase decisions, do not take into account their neighbors' future actions (which might change their external values of the product).

Contributions. Comparing with their positive counterparts, negative externalities possess more irregularity and pose more challenges for research on product diffusion, especially from the perspective of pricing. The intuitive hardness is confirmed by the following theoretical intractability.

- By a reduction from the 3SAT problem we show that finding an optimal iterative pricing is NP-hard even for the extremely simple case of unweighted tree network with uniform intrinsic values.

Complementary to the hardness result,

- We design a 2-approximation algorithm for iterative pricing in general weighted networks with general intrinsic values. An exact $O(n^2)$ -time algorithm is designed for unweighted split networks with uniform intrinsic values.

The 2-approximation algorithm is remarkable for its simplicity and versatility to handle the most general problem regardless of network topologies, link weights or intrinsic values. We also study single pricing as an approximation of iterative pricing, and obtain the following negative and positive results

- We prove that optimal single pricing can be arbitrarily worse (at a rate of $\ln \ln n$) than the optimal iterative pricing; and on the other hand, optimal single pricing provides nice approximations to

the optimal iterative pricing for several well-known unweighted networks with uniform intrinsic values: $(\ln n)$ -approximation for general networks, 1.5-approximation for forest networks, $(1+\epsilon)$ -approximation a.a.s for Erdős-Rényi networks, and 2-approximation a.a.s. for Barabási-Albert networks (a.k.a. preferential attachment networks).

This justifies the importance of the research of both iterative pricing and single pricing, whose relations in various scenarios represent different trade-offs between revenue efficiency and algorithmic simplicity.

Related work. In the economics literature, the importance of network effects and network externalities in business began to attract serious attention around three decades ago ([14, 19]). Under the most popular frameworks, network effects are assumed to be global instead of local. Namely, only complete networks are considered. Consumers may also act sequentially as in this paper, but are usually assumed to be completely rational in the way that they are able to forecast the decisions of later ones and make their purchasing decisions accordingly. There are quite a lot of followups, most of which are beyond the scope of this paper. We refer the reader to [22] for a most recent development in this paradigm with relaxations of assumptions on consumers.

In the literature of computer science, network pricing stems mainly from the study of diffusion and cascading. One of the most important differences between this strand of research and that of economics is arguably that network structures are explicitly and seriously addressed. Over the last decade, under the framework of viral marketing, the algorithmic study of diffusing products with positive externalities is especially fruitful for influence maximization, see, e.g., [11, 20, 21]. To the best of our knowledge, Hartline et al. [17] was the first to study the diffusion problem from a network pricing perspective. They investigated marketing strategies for revenue maximization with positive externalities. Consumers are visited in a sequence (determined by the seller), and asked whether to buy or not under some price (different consumers may receive different prices, referred to as differential pricing or discriminative pricing). They showed that for myopic consumers, a reasonable approximation of the optimal marketing strategy can be achieved in a simple way of influence-and-exploit. While complete information was assumed in [17], Chen et al. [12] studied the incomplete information model with rational players and positive externalities. They provided a polynomial time algorithm that computes all the pessimistic (and optimistic) equilibria and the optimal single price. When discriminative pricing is allowed, they proved the NP-hardness of optimal equilibrium computation, and gave an FPTAS for the case that consumers are already partitioned into groups such that those within the same group must receive the same price.

Iterative pricing, with a very limited literature, was discussed by Akhlaghpour et al. [1] for positive externalities. The authors studied two iterative pricing models in which consumers are assumed to be myopic. In the first model, they gave an FPTAS for the optimal pricing strategy in the general case. In the second model, they showed that the revenue maximization problem is inapproximable even in some special case. Their second model is quite similar to ours.

Although there is also a large literature in the field of classical economics studying negative externalities (under various terms, e.g. the Veblen effect, the snob effect, the congestion effect etc.), explicit networks are rarely treated seriously as aforementioned. One of the classical papers in this strand is [18], where the nuclear weapon selling problem was considered from the perspective of network effects. In the more recent computer science literature, compared with positive externalities, network pricing problems with negative externalities are much less investigated. Chen et al. [12] showed that when both positive and negative externalities are allowed in their model, computing any approximate equilibrium is PPAD-hard. However, the complexity status of the problem in the case with only negative externalities is still unknown. The only paper known to us that deals with the network pricing problem with negative externalities is [5] by Bhattacharya et al., although their main focus is on equilibrium computation for given prices rather than pricing. The authors also considered linear externalities, but a combination of single pricing, complete information and strategic consumers. They showed that for any given price, the game that the consumers play is an exact potential game, and provided a set of hardness results. They proved that finding the best equilibrium is NP-hard even for trees, and gave a 2-approximation algorithm for bipartite networks. Along a different line, Alon et al.

[2] used the term “negative externality” to mean the harm of discriminative pricing on consumers (because discriminative pricing gives many consumers a feeling of inequality).

All the papers cited above assume that externalities are only exerted between consumers who buy the product. In contrast, for some products or services, e.g., public goods, externalities are exerted from purchasers to nonpurchasers. Our paper is close to [9] in the sense that both papers address strategic substitutes (each player has less incentive to buy when more neighbors purchase), although the network externalities are negative in our settings but positive in their settings of public goods. In the computer science, the public goods pricing problem was also studied by Feldman et al. [15]. Their work differs from ours in two main respects: (i) In our externality model, a consumer’s utility is subtractive over the purchases made by this neighbors, whereas in their setting, purchases of neighbors are substitutes. (ii) Technically, they related the pricing problem (where externalities in their model are mathematically expressed in terms of products of neighbors actions) to a single-item auction problem, while we address the pricing problem (where externalities are expressed in terms of sums of neighbors actions) using iterative algorithmic approaches. As noted by the authors [15], their results carry over to a special kind of negative externality, where the valuation of a consumer on the product is positive if and only if the consumer is the only one among her/his neighbors who possess the product. The aforementioned literature are all on indivisible goods. The network pricing problems for divisible goods with quadratic utilities functions have been studied in [6, 10]. Along with [15], a growing number of papers have been addressing the network externality problem from the perspective of mechanism design and auction theory (e.g. [4, 13, 16]).

The remainder of the paper is organized as follows. Section 2 gives the mathematical formulation of our iterative pricing model. Section 3 is devoted to general iterative pricing, including NP-hardness (Section 3.1), 2-approximation for general weighted network with general intrinsic values (Section 3.2) and optimal pricing for unweighted split network with uniform intrinsic values (Section 3.3). Section 4 discusses the relation between single pricing and iterative pricing. Single pricing is shown to guarantee 1.5-approximation for forests (Section 4.1), near optimal for Erdős-Rényi networks (Section 4.2), $(2-\epsilon)$ -approximation for Barabási-Albert networks (Section 4.3), and approximation with ratio within $[\ln \ln n, \ln n]$ for general networks (Section 4.4). Section 5 concludes the paper with remarks on future research.

2 The model

Let $G = (V, E)$ be the given undirected network (without self-loops, and possibly associated with a nonnegative integer weight function $w \in \mathbb{Z}_+^{V \times V}$), where $V \equiv [n]$ is the set of n consumers, and E represents the links between pairs of consumers. When the weight function $w \in \mathbb{Z}_+^{V \times V}$ is discussed, it is always assumed that $w_{ij} = w_{ji}$ for all $i, j \in V$ and $w_{ij} = 0$ if and only if $ij \notin E$. Given any consumer $i \in V$ and subset $S \subseteq V$ of consumers, we use $w_i(S) = \sum_{j \in S} w_{ij}$ to denote the sum of weights contributed to consumer i by those in S . Clearly, only i ’s neighbors can possibly contribute.

We name the model under investigation as PNC (*Pricing with Negative externalities and Complete information*). Let Q , which usually shrinks as the iterative pricing proceeds, denote the set of consumers who do not possess the product. Each consumer $i \in V$ has an intrinsic value $\nu(i) \in \mathbb{R}_+$, and her *total value* of the product equals $\nu(i) + w_i(Q)$. Initially $Q = V$. The PNC model proceeds as follows.

- *Iterative pricing.* The monopolist seller announces prices p_1, p_2, \dots, p_τ sequentially at time $1, 2, \dots, \tau$.
- *Impatient consumers.* As soon as a price is announced, a consumer in Q buys the product if and only if her current total value is greater than or equal to the current price.
- *Simultaneous moves.* We assume that, for each newly announced price, all consumers in Q make their decisions (buying or not buying) simultaneously.

Note that a consumer in Q who does not purchase at current time t under price p_t may be willing to buy at a later time $t' > t$ under a lower price $p_{t'} < p_t$. For each $t = 1, 2, \dots, \tau$, let $B(p_t)$ denote the set of consumers who buy the product at price p_t , (i.e., at time t , or in the t -th *round*). We use $R(\mathbf{p})$ to denote

the revenue derived from $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$, i.e., $R(\mathbf{p}) = \sum_{t=1}^\tau p_t \cdot |B(p_t)|$. In case of $\mathbf{p} = (p_1)$, we often write $R(\mathbf{p})$ as $R(p_1)$. The PNC problem is to find a pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ such that $R(\mathbf{p})$ is maximized, where both the length τ and the entries p_1, p_2, \dots, p_τ of the sequence are variables to be determined.

3 General iterative pricing

In this section, we study the PNC model in the most general setting where no restriction is imposed to the length of the pricing sequence.

3.1 NP-hardness

We prove that finding an optimal pricing sequence for the PNC model is NP-hard, even when the intrinsic values are all zero, link weights are unit, and the network is a tree. Throughout this subsection, we assume that the intrinsic values of all consumers are zero.

We begin with some preliminaries that will be used in the formal proofs. Let $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ be a pricing sequence. For any $t \in [\tau]$ and $i \in V$, we use $\nu_t(i, \mathbf{p})$ to denote the (total) value of the product at time t (in the t -th round) for consumer i during the selling/purchase process. Since intrinsic value $\nu(i)$ is zero by assumption, $\nu_t(i, \mathbf{p})$ is the sum of weights from i 's neighbors who have not purchased yet in the previous rounds.

Observation 3.1. During the selling process, the value of the product for each consumer i does not increase, i.e. $\nu_{t+1}(i, \mathbf{p}) \leq \nu_t(i, \mathbf{p})$ for all $t = 1, 2, \dots, \tau - 1$.

Given a subset of nodes $S \subseteq V$, we use $R_{t,S}(\mathbf{p})$ to denote the revenue from these consumers until time t , i.e., $R_{t,S}(\mathbf{p}) = \sum_{i=1}^t |S \cap B(p_i)| p_i$. For brevity, we also write $R_{\tau,S}(\mathbf{p})$ as $R_S(\mathbf{p})$. In particular, we have $R_V(\mathbf{p}) = R(\mathbf{p})$.

Definition 3.2. We call pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ *irredundant* if for each $i \in [\tau]$, there is at least one consumer who purchases under price p_i .

Every pricing sequence \mathbf{p} is “equivalent” to a unique irredundant pricing sequence \mathbf{p}' which is derived from \mathbf{p} by removing all prices under which no consumers purchase. Clearly, the equivalent pricing sequences bring about the same revenue $R(\mathbf{p}) = R(\mathbf{p}')$. This allows us to focus on irredundant pricing sequences.

Observation 3.3. If pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ is irredundant, then it is decreasing, i.e. $p_1 > p_2 > \dots > p_\tau$.

Assume that $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ is an irredundant pricing sequence. Since by Observation 3.3 the entries of \mathbf{p} are all distinct, we also view \mathbf{p} as a set $\{p_1, p_2, \dots, p_\tau\}$, and use the symbol $p_i \in \mathbf{p}$ to mean that p_i is an entry of the pricing sequence \mathbf{p} . For all $1 \leq t \leq \tau$, define $B_t(\mathbf{p}) = \cup_{i=1}^t B(p_i)$ to be the set of consumers who have purchased in the first t rounds. For notational convenience, we set $B_0(\mathbf{p}) = B(p_0) = \emptyset$. Recall that in the PNC model we have assumed that consumers are all impatient in the sense that they will definitely purchase as long as the current price is lower than or equal to their current values. As \mathbf{p} is irredundant, $B(p_1), B(p_2), \dots, B(p_\tau)$ can be computed in a recursive way: $B(p_t) = \{i \in V : w_i(V \setminus B_{t-1}(\mathbf{p})) \geq p_t\}$, $t = 1, 2, \dots, \tau$.

Definition 3.4. A pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ is called *normal* if it is irredundant, and for any $i \in [\tau]$ and any $\epsilon > 0$, increasing p_i to $p_i + \epsilon$ (other prices remain the same) changes the set of consumers who purchase at the i -th round.

Clearly, all entries of a normal pricing sequence are integers. Given an irredundant pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ together with $B(p_1), B(p_2), \dots, B(p_\tau)$, one can easily compute a normal pricing sequence $\mathbf{p}' = (p'_1, p'_2, \dots, p'_\tau)$, which is “equivalent” to \mathbf{p} in the sense that $B(p'_t) = B(p_t)$ for all $t \in [\tau]$, as follows: $p'_t = \min\{w_i(V \setminus B_{t-1}(\mathbf{p})) : i \in B(p_t)\}$, $t = 1, 2, \dots, \tau$. It is clear that $R(\mathbf{p}') \geq R(\mathbf{p})$. The following observation enables us to concentrate on normal pricing sequences in our NP-hardness proofs.

Observation 3.5. There is an optimal pricing sequence that is normal.

The NP-hardness for the PNC model is proved by reduction from the 3SAT problem. The input of the 3SAT problem are n boolean variables x_1, x_2, \dots, x_n , and m clauses $c^j = (x^{j1} \vee x^{j2} \vee x^{j3})$, $1 \leq j \leq m$, where $x^{j\ell}$ is a literal taken from $\{x_1, x_2, \dots, x_n, \neg x_1, \neg x_2, \dots, \neg x_n\}$, $1 \leq j \leq m, 1 \leq \ell \leq 3$. For convenience, we write $x^{j\ell} \in c^j$. The 3SAT problem is to determine if there is a satisfactory truth assignment to the n variables that makes all m clauses evaluate to TRUE. To avoid triviality, we assume that $m \geq 3$, and for each $i \in [n]$, there exist $j, j' \in [m]$ such that $x_i \in c^j$ and $\neg x_i \in c^{j'}$.

Next, we prove the NP-hardness for the weighted case with general network structures. The proof, which highlights the high level idea in our later proof to handle the unweighted case with tree structures, turns out to be much easier to understand.

Theorem 3.6. *In the PNC model, computing an optimal pricing sequence is NP-hard, even when all the intrinsic values are zero.*

Proof. Our reduction here uses a slightly restricted version of the 3SAT problem, the 3-OCC-3SAT problem, which is known to be NP-hard, where for each $i \in [n]$, there are at most three clauses that contains either x_i or $\neg x_i$.

For any instance I of the 3-OCC-3SAT problem, we construct an instance P of the network pricing problem on network $G = (V, E)$ as follows. There are a total of $5n + 3m$ nodes:

- For each variable x_i , there is a gadget V_i . Each pair of literals x_i and $\neg x_i$ are simulated by two nodes (with the names unchanged), respectively, and three auxiliary ones, y_{i1}, y_{i2}, y_{i3} . See the left part of Figure 1 for the links and their weights.
- For each clause $c^j = (x^{j1} \vee x^{j2} \vee x^{j3})$, there is a gadget C^j . The clause is simulated by a node c^j and two auxiliary ones, d^j and e^j . See the right part of Figure 1 for the links and weights among them.
- A literal node (x_i or $\neg x_i$) is linked to a clause node c^j if and only if this literal appears in the clause, and the weight of the link is 1.
- The integer parameters in the weights satisfy

$$a_1 > 5a_2 > \dots > 5^{i-1}a_i > \dots > 5^{n-1}a_n > 5^n a > 5^{n+1}mn. \quad (3.1)$$

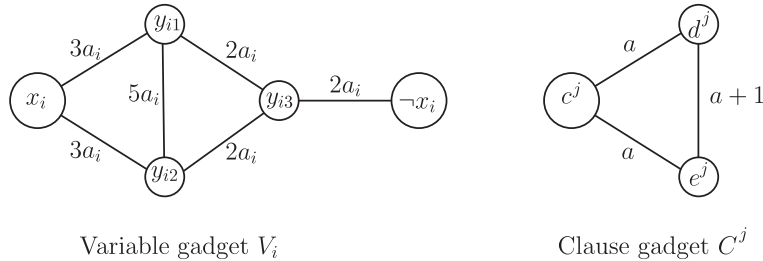


Figure 1: Literal nodes and clause nodes are represented by larger circles, while auxiliary nodes are represented by smaller ones.

Obviously, the above construction can be done in polynomial time. Observe first that all the consumers in the variable gadgets are incident with links of weights much larger than the total weight of links that are incident with any clause consumer. This structure permits us to consider the variable consumers before the clause ones. In the rest of this proof, we may abuse the notations V_i and C^j a little bit to represent both the gadgets and the corresponding node sets, respectively.

Due to Observation 3.5, we only consider normal pricing sequences. Given any normal pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$, let ξ be the first time that the price is equal to or lower than $2a + 3$, i.e.,

$$\xi = \min\{t : p_t \in \mathbf{p}, p_t \leq 2a + 3\}.$$

Note that before time ξ , no consumer in the clause gadgets has purchased, i.e.,

$$B_{\xi-1}(\mathbf{p}) \cap (\cup_{j=1}^m C^j) = \emptyset.$$

The key idea of our proof is simple: we shall show that for each pair of nodes x_i and $\neg x_i$, we can sell the product to one and only one of them, and this makes no difference for the revenue at all before time ξ (see Claim 1 below). The only difference that the choice between x_i and $\neg x_i$ makes is upon the clause gadget nodes after time ξ . Our construction makes these choices really hard because they correspond to a (possible) solution of the 3-OCC-3SAT problem.

For any $i \in [n]$, we note that $w_v(V) \leq 10a_i$ for all $v \in V_i$; thus no consumer in V_i purchases when the price is above $10a_i$.

Claim 1. For all $i = 1, 2, \dots, n$,

- (i) if $\mathbf{p} \cap [2a_i, 10a_i] = \{10a_i, 2a_i\}$, then $R_{\xi-1, V_i}(\mathbf{p}) = 24a_i$, $x_i \notin B_{\xi-1}(\mathbf{p})$ and $\neg x_i \in B_{\xi-1}(\mathbf{p})$;
- (ii) if $\mathbf{p} \cap [2a_i, 10a_i] = \{6a_i\}$, then $R_{\xi-1, V_i}(\mathbf{p}) = 24a_i$, $x_i \in B_{\xi-1}(\mathbf{p})$ and $\neg x_i \notin B_{\xi-1}(\mathbf{p})$;
- (iii) $R_{\xi-1, V_i}(\mathbf{p}) \leq 24a_i$, and the equality holds if and only if $\mathbf{p} \cap [2a_i, 10a_i] \in \{\{10a_i, 2a_i\}, \{6a_i\}\}$.

Statements (i) and (ii) are easily checked. It remains to prove $R_{\xi-1, V_i}(\mathbf{p}) < 24a_i$ if $\mathbf{p} \cap [2a_i, 10a_i] \notin \{\{10a_i, 2a_i\}, \{6a_i\}\}$. Note that $R_{\xi-1, V_i}(\mathbf{p}) < 24a_i$ is trivial if $\mathbf{p} \cap [2a_i, 10a_i] = \emptyset$. Hence we may assume that there exists a maximum price $\hat{p} \in \mathbf{p} \cap [2a_i, 10a_i]$. By normality of \mathbf{p} , we know that $\hat{p} \in \{10a_i, 6a_i + h_i, 6a_i, 2a_i + h'_i, 2a_i\}$, where h_i and h'_i are total weights that x_i and $\neg x_i$ get from clause gadgets, respectively. Hence $h_i + h'_i \leq 3$ (recall the definition of 3-OCC-3SAT). For the case that $\hat{p} \leq 2a_i + h'_i$, it is obvious that $R_{\xi-1, V_i}(\mathbf{p}) < 24a_i$. We are left to the analysis of the remaining three cases, which will establish Statement (iii).

- $\hat{p} = 10a_i$. It follows from $\mathbf{p} \cap [2a_i, 10a_i] \neq \{2a_i, 10a_i\}$ that $\mathbf{p} \cap [2a_i, 10a_i] = \{10a_i\}$, because the only price in $[2a_i, 10a_i]$ that is smaller than $10a_i$ and makes \mathbf{p} normal is $2a_i$. This gives $R_{\xi-1, V_i}(\mathbf{p}) = 20a_i < 24a_i$.
- $\hat{p} = 6a_i + h_i$. The normality of \mathbf{p} implies $\mathbf{p} \cap [2a_i, 10a_i] \in \{\{6a_i + h_i, 2a_i + h'_i\}, \{6a_i + h_i, 2a_i\}, \{6a_i + h_i\}\}$ and hence $R_{\xi-1, V_i}(\mathbf{p}) \leq 3(6a_i + h_i) + 4a_i < 24a_i$.
- $\hat{p} = 6a_i$. An argument similar to the previous case shows that $R_{\xi-1, V_i}(\mathbf{p}) \leq 3 \times 6a_i + 4a_i < 24a_i$. ■

Claim 2. For each $1 \leq j \leq m$, $R_{C^j}(\mathbf{p}) \leq 6a + 3$, and the equality holds if and only if $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap [2a, 2a + 3] = \{2a + 1\}$, where $D^j = \{x^{j1}, x^{j2}, x^{j3}\}$.

It is easy to check that when $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap [2a, 2a + 3] = \{2a + 1\}$, the equality $R_{C^j}(\mathbf{p}) = 6a + 3$ holds. We prove $R_{C^j}(\mathbf{p}) < 6a + 3$ in the other cases. When $D^j \subseteq B_{\xi-1}(\mathbf{p})$ or $\mathbf{p} \cap (2a, 2a + 3] = \emptyset$, it is easy to see that $R_{C^j}(\mathbf{p})$ is at most $2a \times 3 = 6a$. So we only need to discuss the case of $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap (2a, 2a + 3] \neq \emptyset$.

Let $\hat{p} \in \mathbf{p} \cap (2a, 2a + 3]$ be maximum. By normality of \mathbf{p} , we know that $\hat{p} \in \{2a + 1, 2a + 2, 2a + 3\}$. Since $w(d^j) = w(e^j) = 2a + 1$, it can be seen that $R_{C^j}(\mathbf{p}) < 6a + 3$ holds for $\hat{p} \in \{2a, 2a + 2, 2a + 3\}$. So Claim 2 is valid. ■

We are now ready to prove the close relation between the 3-OCC-3SAT instance I and the PNC instance P . Let $\text{OPT}(P)$ be the optimal objective value of P . Define

$$L = \sum_{i=1}^n 24a_i + m(6a + 3).$$

Claim 3. $\text{OPT}(P) \leq L$.

Suppose that \mathbf{p} is an optimal solution of P . We can assume without loss of generality that $p_\tau > 0$. Let $U = V \setminus B_\tau(\mathbf{p})$ denote the set of consumers who do not purchase during the whole selling process. Note first from our previous discussion that $R_{V_i}(\mathbf{p})$ may be greater than $24a_i$, although $R_{C^j}(\mathbf{p}) \leq 6a + 3$ holds for every $j \in [m]$. However, $R_{V_i}(\mathbf{p}) \leq 24a_i + 3$ is always valid, because $\sum_{v \in V_i} w_v(\cup_{j=1}^m C^j) \leq 3$ (recall the definition of 3-OCC-3SAT and the construction of P). Also, when $V_i \setminus U \subseteq B_{\xi-1}(\mathbf{p})$, we do have $R_{V_i}(\mathbf{p}) \leq 24a_i$. Therefore, if $(\cup_{i=1}^n V_i) \setminus U \subseteq B_{\xi-1}(\mathbf{p})$, the above claim is derived immediately from Claims 1 and 2.

Suppose some $v \in V_i \setminus U$ purchases at price $p_t > 0$ with $t \geq \xi$. If $v \in \{y_{i1}, y_{i2}, y_{i3}\}$, or $v \in \{x_i, \neg x_i\}$ and $p_t > 3$, then it can be seen easily that $R_{\xi-1, V_i}(\mathbf{p}) \leq 20a_i$, and hence (3.1) implies $R_{V_i}(\mathbf{p}) \leq 20a_i + 5(2a + 3) < 20a_i + 2a_i + 15 < 24a_i$. It remains to consider the case where $\{y_{i1}, y_{i2}, y_{i3}\} \subseteq B_{\xi-1}(\mathbf{p})$, $0 < p_t \leq 3$, and $v \in \{x_i, \neg x_i\} \cap D^{j_0}$ for some $j_0 \in [m]$ with $c^{j_0} \notin B_{t-1}(\mathbf{p})$. Since c^{j_0} does not purchase before time t , it must be the case that $\mathbf{p} \cap (p_t, 2a + 1] = \emptyset$. It follows $p_t \leq 3$ that $R_{C^j}(\mathbf{p}) \leq 9$ for all $j \in [m]$. Hence (3.1) implies $R(\mathbf{p}) = R_V(\mathbf{p}) \leq \sum_{i=1}^n (24a_i + 3) + 9m < L$. So Claim 3 is indeed correct. \blacksquare

To establish the NP-hardness of the pricing problem, it suffices to prove that

$$\text{OPT}(P) \geq L \Leftrightarrow I \text{ is satisfiable.}$$

(\Leftarrow) Suppose that I has a satisfactory truth assignment π with s variables assigned “TRUE” and the remaining $n - s$ variables assigned “FALSE”. Let pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_{n+s+1})$ be a solution to P such that

- There are one or two prices for each variable gadget depending on whether the variable is assigned “TRUE” or “FALSE” in π : if x_i is assigned “TRUE” then $10a_i, 2a_i \in \mathbf{p}$, if x_i is assigned “FALSE” then $6a_i \in \mathbf{p}$;
- There is a common price for the m clause gadgets: $p_{n+s+1} = 2a + 1 \in \mathbf{p}$.

For this \mathbf{p} , note from (3.1) that $\xi = n + s + 1$. According to Claim 1, $R_{\xi-1, V_i}(\mathbf{p}) = 24a_i$ for each $i \in [n]$. For each clause gadget C^j , due to Claim 1(i) and (ii), we know that consumer $x^{j\ell} \in B_{\xi-1}(\mathbf{p})$, $1 \leq \ell \leq 3$, if and only if the corresponding literal is “FALSE” in π . Since π is a satisfactory assignment, we know that there is at least one literal in x^{j1}, x^{j2}, x^{j3} that is assigned true. Therefore, for each j , it holds that $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$. Combining this fact with Claim 2, we know that $R_{\xi, C^j}(\mathbf{p}) = 6a + 3$ for each clause C^j . This completes the sufficiency part.

(\Rightarrow) Suppose now $\text{OPT}(P) \geq L$. Due to Observation 3.5, there exists a normal pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ whose objective value $R(\mathbf{p})$ is at least L . Combining with Claim 3, this can only be the case that $\text{OPT}(P) = L$. By arguments in the proof of Claim 3, we know that conditions in Claims 1 and 2 must hold. We construct a truth assignment π as follows: for each $i \in [n]$, if $\mathbf{p} \cap [2a_i, 10a_i] = \{2a_i, 10a_i\}$, we assign “TRUE” to variable x_i . Otherwise, that is $\mathbf{p} \cap [2a_i, 10a_i] = \{6a_i\}$, we assign “FALSE” to x_i . By Claim 2, we know that $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ for all $j \in [m]$. Therefore π is indeed a satisfactory truth assignment for I . This completes the necessity part, and therefore the proof of Theorem 3.6. \square

A corollary of the above proof says that the length of the optimal pricing sequence of the PNC problem can not be upper bounded by any constant. This remains true for the unweighted trees without intrinsic values, as seen from the proof of the following stronger NP-hardness result.

Theorem 3.7. *In the PNC model, computing an optimal pricing sequence is NP-hard, even when the underlying network is an unweighted tree and all the intrinsic values are zero.* \square

While the proof, which we postpone to the appendix, has high level similarities to the one for Theorem 3.6, a substantially more careful approach is required to handle the acyclic structure, and new ideas are needed to simulate the weights with unweighted links.

In view of the above NP-hardness result, it is desirable to design good approximation algorithms for the general PNC problem and exact algorithms for special cases. In the following, we obtain 2-approximation for the general case (Theorem 3.8), and an optimal pricing for unweighted split networks (Theorem 3.9).

3.2 2-approximation

As to approximation, we find that, more or less surprisingly, a very simple greedy algorithm performs fairly well, achieving 2-approximation for the most general scenario.

For any subnetwork H of G , and any $i \in V(H)$, where $V(H)$ is the node set of H , we use $d_H^w(i) = \sum_{j \in V(H)} w_{ij}$ to denote the weighted degree of i in H . For any real function f and any nonempty subset S of its domain, let $f(S) = \sum_{s \in S} f(s)$.

ALGORITHM 1: Iterative Pricing

Input: Network $G = (V, E)$ with weight function $w \in \mathbb{Z}_+^{V \times V}$ and intrinsic value function $\nu \in \mathbb{R}_+^V$.

Output: Sequence \mathbf{p} of prices.

1. $G_0 \leftarrow G, \quad t \leftarrow 0$
 2. **While** $V(G_t) \neq \emptyset$ **do**
 3. $t \leftarrow t + 1$
 4. $p_t \leftarrow \max\{\nu(i) + d_{G_{t-1}}^w(i) : i \in V(G_{t-1})\}$
 5. $G_t \leftarrow G_{t-1} \setminus B(p_t)$
 6. **End-while**
 7. Output $\mathbf{p} \leftarrow (p_1, p_2, \dots, p_t)$
-

Theorem 3.8. *For the PNC model, Algorithm 1 finds a 2-approximate pricing sequence in $O(n^2)$ time.*

Proof. Let \mathbf{p}^* be an optimal pricing, and $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ be the pricing output by the algorithm. Since each link $ij \in E$ can contribute at most $2w_{ij}$ to consumers' total values (w_{ij} to each of i and j), we see that

$$R(\mathbf{p}^*) \leq \nu(V) + 2w(E).$$

On the other hand, the definition of p_t in Step 4 of Algorithm 1 guarantees that $B(p_t) = \operatorname{argmax}\{\nu(i) + d_{G_{t-1}}^w(i) : i \in V(G_{t-1})\}$ and therefore each link $ij \in E(G_{t-1})$ with $i \in B(p_t)$ contributes w_{ij} to i 's total value, giving

$$p_t \cdot |B(p_t)| = \nu(B(p_t)) + \sum_{i \in B(p_t)} d_{G_{t-1}}^w(i) \geq \nu(B(p_t)) + w(E(G_{t-1}) - E(G_t)), \text{ for all } t \in [\tau].$$

It follows that

$$\begin{aligned} R(\mathbf{p}) &= \sum_{t=1}^{\tau} p_t \cdot |B(p_t)| \geq \sum_{t=1}^{\tau} \nu(B(p_t)) + \sum_{t=1}^{\tau} w(E(G_{t-1}) - E(G_t)) \\ &= \nu(\cup_{t=1}^{\tau} B(p_t)) + \sum_{t=1}^{\tau} w(E(G_{t-1})) - w(E(G_t)) \\ &= \nu(V) + w(E(G_0)) - w(E(G_\tau)) \\ &= \nu(V) + w(E), \end{aligned}$$

where $G_0 = G$ and $G_\tau = \emptyset$ are guaranteed by Steps 1 and 2. Hence

$$R(\mathbf{p}^*)/R(\mathbf{p}) \leq (\nu(V) + 2w(E))/(\nu(V) + w(E)) \leq 2$$

justifies the approximation ratio 2.

To see the $O(n^2)$ running time, we note that the while-loop repeats $\tau \leq n$ times, and each repetition finishes in $O(n)$ time. \square

3.3 Optimal pricing for unweighted split networks

Network $G = (V, E)$ is a *split network* if its node set V can be partitioned into two sets C and I such that C induces a clique and I is an independent set of G . Clearly, the nodes in I can only have neighbors in C . In case of each node in I adjacent to exactly one node in C , network G is called *core-peripheral*. Core-peripheral networks are widely accepted as good simplifications of many real-world networks and thus have been extensively studied in various environments [8].

We consider the case of uniform intrinsic values, which can be assumed w.l.o.g. to be zeros. Let $d(v) = d_G(v)$ denote the degree of $v \in V$ in G . Suppose that $C = \{v_1, v_2, \dots, v_k\}$, and $d(v_i) \leq d(v_{i+1})$ for every $i \in [k-1]$. For each $i \in [k]$, note that v_1, \dots, v_i form a clique set C_i and their neighbors in I form an independent set I_i , and $C_i \cup I_i$ induces a split subnetwork G_i of G with degree sequence

$$d_{G_i}(u_{\ell_i}^i) \leq d_{G_i}(u_{\ell_i-1}^i) \leq \dots \leq d_{G_i}(u_1^i) \leq d_{G_i}(v_1) \leq \dots \leq d_{G_i}(v_i),$$

where $I_i = \{u_1^i, u_2^i, \dots, u_{\ell_i}^i\}$. Apparently, $d_{G_i}(v_h) = d(v_h) - (k - i)$ for every $h \in [i]$. Consider an optimal pricing $\mathbf{p} = (p_1, \dots, p_\tau)$ for the PNC problem on G_i , and write the corresponding maximum revenue as $\text{OPT}(G_i)$. One of the following must hold.

- $p_1 = d_{G_i}(v_{h+1})$ for some $h \in [i-1]$, and exactly $(i-h)$ nodes, i.e., v_{h+1}, \dots, v_i , purchase at price p_1 , offering revenue $(i-h)p_1 = (i-h)d_{G_i}(v_{h+1})$. It follows that $\tau \geq 2$ and (p_2, \dots, p_τ) is an optimal pricing for G_h , giving $\text{OPT}(G_i) = (i-h)d_{G_i}(v_{h+1}) + \text{OPT}(G_h) = (i-h)d(v_{h+1} - k + i) + \text{OPT}(G_h)$.
- $p_1 = d_{G_i}(u_j^i)$ for some $j \in [\ell_i]$ and exactly $(i+j)$ nodes, i.e., $u_j^i, u_{j-1}^i, \dots, u_1^i, v_1, v_2, \dots, v_i$, purchase at price p_1 , offering revenue is $(i+j)p_1 = (i+j)d_{G_i}(u_j^i)$. Since the nodes not purchasing at price p_1 are pairwise nonadjacent, it is easy to see that $\mathbf{p} = (p_1)$ and $\text{OPT}(G_i) = (i+j) \cdot d_{G_i}(u_j^i)$.

For convenience, let $\text{OPT}(G_0)$ stands for real number 0. Then $\text{OPT}(G) = \text{OPT}(G_k)$ can be computed by the following recursive formula:

$$\text{OPT}(G_i) = \max \left\{ \max_{h=0}^{i-1} \{ \text{OPT}(G_h) + (d(v_{h+1}) - k + i)(i-h) \}, \max_{j=1}^{\ell_i} \{ (i+j) \cdot d_{G_i}(u_j^i) \} \right\} \text{ for } i = 1, 2, \dots, k.$$

This formula implies the following result.

Theorem 3.9. *For the PNC model, an optimal pricing sequence for any unweighted split network with uniform intrinsic values can be found in $O(n^2)$ time by dynamic programming.* \square

4 Approximation by single pricing

Finding an optimal single pricing is trivial because it can be chosen from the n total values of the consumers. Thus it is natural to ask: How does the optimal single pricing work as an approximation to the optimal iterative pricing? We find that the answer is both “good” and “bad”, in the sense that single pricing works rather well for many interesting networks with unit weights and uniform intrinsic values, including forests, Erdős-Rényi networks and Barabási-Albert networks, but in general, its worst-case performance, even when restricted to unweighted networks, can be arbitrarily bad. This justifies the importance of the research of iterative pricing, and at the same time poses the interesting question of investigating the relation between single pricing and iterative pricing for more realistic scenarios.

In this section, we restrict our attention to unweighted networks G with uniform intrinsic values, for which we may assume without loss of generality that all intrinsic values are zero, and use $\text{OPT}(G)$ to denote the revenue derived from an optimal iterative pricing.

4.1 1.5-approximation for forests

We show that the best single price guarantees an approximate ratio of 1.5 for unweighted forests with uniform intrinsic values.

Theorem 4.1. *For the PNC model, the single pricing p with maximum $p \cdot |B(p)|$ has an approximation ratio of 1.5 for unweighted forests with uniform intrinsic values.*

Proof. Suppose that forest $G = G_0$ consists of k components (trees) $G_h = (V_h, E_h)$, $h = 1, \dots, k$. Let ℓ_h denote the number of leaves in G_h . Note that $\sum_{i=1}^k \text{OPT}(G_i) \geq \text{OPT}(G_0)$, and

$$p \cdot |B(p)| \geq \max\{|B(1)|, 2|B(2)|\} = \max\{|V_0|, 2(|V_0| - \ell_0)\} \geq \frac{2}{3} (2|V_0| - \ell_0) = \frac{2}{3} \left(2 \sum_{i=1}^k |V_i| - \sum_{i=1}^k \ell_i \right).$$

It suffices to show that $\text{OPT}(G_i) \leq 2|V_i| - \ell_i$ for each $i \in [k]$, in order to guarantee that the approximation is at most 1.5.

If G_i is a star network or a link, then $\text{OPT}(G_i) = |V_i| \leq 2|V_i| - \ell_i$. Suppose that G_i is neither a star nor a link, and let G'_i be the tree obtained from G_i by deleting all its leaves. Clearly, $d_{G'_i}(v) \leq d_{G_i}(v)$ for every non-leaf node of G'_i .

Let \mathbf{p} be an optimal pricing for G_i . Consider an arbitrary leaf u of G'_i . Let $L(u)$ denote the set of u 's leaf neighbors in G_i . Under \mathbf{p} , either u purchases before all nodes in $L(u)$ at a price higher than 1 or all nodes in $\{u\} \cup L(u)$ purchase at price 1. As u has at least one non-leaf neighbor in G_i , it is easy to see that in either case, the total payment by nodes in $\{u\} \cup L(u)$ is upper bounded by $d_{G_i}(u) \leq d_{G'_i}(u) + |L(u)|$. Hence

$$\begin{aligned} \text{OPT}(G_i) &\leq \sum_{\text{non-leaf node } v \text{ of } G'_i} d_{G_i}(v) + \sum_{\text{leaf node } u \text{ of } G'_i} (d_{G'_i}(u) + |L(u)|) \\ &= \sum_{\text{non-leaf node } v \text{ of } G'_i} d_{G'_i}(v) + \sum_{\text{leaf node } u \text{ of } G'_i} (d_{G'_i}(u) + |L(u)|) \\ &= \ell_i + \sum_{v \in V_i} d_{G'_i}(v) \\ &= \ell_i + 2(|V_i| - \ell_i - 1) \\ &< 2|V_i| - \ell_i, \end{aligned}$$

as desired. □

Remark 4.2. In Theorem 4.1, to achieve the approximation ratio 1.5, the single price can be simply chosen between 1 and 2, whichever produces a larger revenue. Moreover, the ratio 1.5 is tight, as shown by the following tree G .

Tree G with $n = 1 + 2k$ nodes is a spider with center of degree k and each leg of length 2 (i.e., the tree obtain from star $K_{1,k}$ by subdividing each link with a node). It is easy to see that the maximum revenue $3k$ is given by pricing sequence $(k, 1)$. However, any single pricing can produce a revenue of at most $\max\{k \cdot 1, 2 \cdot (k+1), 1 \cdot (2k+1)\} = 2k+2$. The tightness follows from $3k/(2k+2) \rightarrow 1.5$ ($k \rightarrow \infty$).

4.2 Near optimal pricing for Erdős-Rényi networks

For large n , there is a simple algorithm that is “almost optimal” for “almost all” Erdős-Rényi networks $\mathbb{G}(n, \eta(n))$. The network is constructed by connecting n nodes randomly; each link is included in the network with probability $\eta(n)$. This algorithm, which will be referred to as $A(\delta)$, prices only once with price $(1 - \delta)(n - 1)\eta(n)$, where $\delta > 0$ is a parameter to be determined by the approximation ratio that we intend to reach.

Theorem 4.3. *Given arbitrarily small positive number $\epsilon > 0$, set $\delta \in (0, 1)$ such that*

$$\frac{1 + \delta}{1 - \delta} < 1 + \epsilon. \tag{4.1}$$

Then for the PNC model, Algorithm A(δ) has an approximation ratio of at most $1 + \epsilon$ for asymptotically almost all networks $\mathbb{G}(n, \eta(n))$, as long as

$$\frac{\eta(n)}{\sqrt{(\ln n)/n}} \rightarrow +\infty. \quad (4.2)$$

To be precise, under condition (4.2), we have

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{2|E(\mathbb{G}(n, \eta(n)))|}{r(\mathbb{G}(n, \eta(n)))} \leq 1 + \epsilon \right) = 1, \quad (4.3)$$

where $E(\mathbb{G}(n, \eta(n)))$ is the link set of $\mathbb{G}(n, \eta(n))$, $\Pr(\cdot)$ is the probability function, and $r(\mathbb{G}(n, \eta(n)))$ is the revenue obtained from the single pricing $(1 - \delta)(n - 1)\eta(n)$.

Proof. Let d_i be the degree of node i in the random network $\mathbb{G}(n, \eta(n))$. As $0 < \delta < 1$, the following Chernoff bound holds:

$$\Pr(|d_i - (n - 1)\eta(n)| > \delta(n - 1)\eta(n)) \leq 2 \exp \left(-\frac{\delta^2(\eta(n))^2(n - 1)}{2} \right). \quad (4.4)$$

Let α_n be the number of nodes in $\mathbb{G}(n, \eta(n))$ whose degrees fall into $[(1 - \delta)(n - 1)\eta(n), (1 + \delta)(n - 1)\eta(n)]$. That is, if we let I_i be an indicator random variable such that $I_i = 1$ if $d_i \in [(1 - \delta)(n - 1)\eta(n), (1 + \delta)(n - 1)\eta(n)]$ and $I_i = 0$ otherwise, then

$$\alpha_n = \sum_{i=1}^n I_i.$$

Now, we use α_n to bound $|E(\mathbb{G}(n, \eta(n)))|$ and $r(\mathbb{G}(n, \eta(n)))$ as follows:

$$\begin{aligned} 2|E(\mathbb{G}(n, \eta(n)))| &\leq \alpha_n(1 + \delta)(n - 1)\eta(n) + (n - \alpha_n)(n - 1), \\ r(\mathbb{G}(n, \eta(n))) &\geq \alpha_n(1 - \delta)(n - 1)\eta(n). \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{2|E(\mathbb{G}(n, \eta(n)))|}{r(\mathbb{G}(n, \eta(n)))} &\leq \frac{\alpha_n(1 + \delta)\eta(n) + (n - \alpha_n)}{\alpha_n(1 - \delta)\eta(n)}, \\ \Pr \left(\frac{2|E(\mathbb{G}(n, \eta(n)))|}{r(\mathbb{G}(n, \eta(n)))} \leq 1 + \epsilon \right) &\geq \Pr \left(\frac{\alpha_n(1 + \delta)p + (n - \alpha_n)}{\alpha_n(1 - \delta)p} \leq 1 + \epsilon \right) \\ &= \Pr(\alpha_n \geq n\epsilon_0), \end{aligned}$$

where $\epsilon_0 = 1/(\epsilon(1 - \delta)\eta(n) - 2\delta\eta(n) + 1)$, which is smaller than 1 due to (4.1). Using $\alpha_n = \sum_{i=1}^n I_i$, we have

$$\begin{aligned} \Pr \left(\frac{2|E(\mathbb{G}(n, \eta(n)))|}{r(\mathbb{G}(n, \eta(n)))} \leq 1 + \epsilon \right) &\geq \Pr \left(\sum_{i=1}^n (I_i - \epsilon_0) \geq 0 \right) \\ &= 1 - \Pr \left(\sum_{i=1}^n (I_i - \epsilon_0) < 0 \right) \\ &\geq 1 - \Pr(I_i - \epsilon_0 < 0 \text{ holds for at some } i \in [n]) \\ &\geq 1 - \sum_{i=1}^n \Pr(I_i - \epsilon_0 < 0) \\ &= 1 - \sum_{i=1}^n \Pr(I_i = 0) \\ &= 1 - \sum_{i=1}^n \Pr(|d_i - (n - 1)\eta(n)| > \delta(n - 1)\eta(n)), \end{aligned}$$

where the second last equality is due to the fact that $I_i \in \{0, 1\}$. It follows from (4.4) and (4.2) that

$$\Pr\left(\frac{2|E(\mathbb{G}(n, \eta(n)))|}{r(\mathbb{G}(n, \eta(n)))} \leq 1 + \epsilon\right) \geq 1 - \sum_{i=1}^n 2 \exp\left(-\frac{\delta^2 p^2(n)(n-1)}{2}\right) \rightarrow 1 \quad (n \rightarrow \infty).$$

This completes the proof. \square

4.3 $(2 - \epsilon)$ -approximation for Barabási-Albert networks

The scale-free property (the power-law tail) has been nicely emulated by the multiple-destination preferential attachment growth model introduced by Barabási and Albert [3]. Starting with a small number of nodes (who are originally connected with each other), at each time step a new node enters network $G = (V, E)$, and attaches to β existing nodes. Each of the existing nodes is attached to the new one with a probability that is proportional to its current degree. Such a process is well-known as the *preferential attachment*. Recall that $|V| = n$. Let $\alpha_{n,k}$ be the fraction of nodes with degree k . It is known from [7] that for any fixed $\epsilon > 0$, and any $\beta \leq k \leq n^{1/15}$,

$$\lim_{n \rightarrow \infty} \Pr\left((1 - \epsilon) \frac{2\beta(\beta + 1)}{k(k + 1)(k + 2)} \leq \alpha_{n,k} \leq (1 + \epsilon) \frac{2\beta(\beta + 1)}{k(k + 1)(k + 2)}\right) = 1. \quad (4.5)$$

Note by the construction that each node has a degree of at least β . Let Γ be the set of all nodes that have a degree of exactly β . Then

$$\Gamma \text{ is an independent set of } G, \quad (4.6)$$

because whenever two nodes are connected, the “older” one must have a degree at least $\beta + 1$. Note also that for any fixed $\epsilon > 0$, the inequality $|E| \leq (1 + \epsilon/2)n\beta$ holds for big enough n .

Theorem 4.4. *Consider the PNC model. For any fixed $\epsilon > 0$, with probability tending to one as $n \rightarrow \infty$, the single pricing with price β achieves an approximation ratio of $2 - 2/(2 + \beta) + \epsilon$ for Barabási-Albert network G . To be precise,*

$$\lim_{n \rightarrow \infty} \Pr\left(\frac{\text{OPT}(G)}{n\beta} \leq 2 - \frac{2}{(2 + \beta)} + \epsilon\right) = 1,$$

where $n\beta$ is the revenue obtained by single price β .

Proof. Given an optimal pricing sequence \mathbf{p} for $G = (V, E)$, we construct a charge c on E as follows: At the time a node $u \in V$ purchases with price p , it must have at least p neighbors, say v_1, \dots, v_p , who have not purchased. We charge each link uv_i ($1 \leq i \leq p$) with 1. After the charge operation is conducted for all nodes, each link $e \in E$ is charged at most twice (i.e. receives charge at most 2). Define $c(e) = 0$ if e is not charged, $c(e) = 1$ if e is charged once, and $c(e) = 2$ if e is charged twice. Note that $c(e) = 2$ only if the both ends of e purchase at the same time (under the same price). The charge function $c : E \rightarrow \{0, 1, 2\}$ satisfies the property that $c(E) = \sum_{p \in \mathbf{p}} p|B(p)|$. For $i = 1, 2$, let C_i consist of links $e \in E$ with $c(e) = i$.

Recall the definition of Γ given above (4.6). We denote by $\delta(\Gamma)$ the set of links that are covered by Γ , and S the set of nodes dominated by Γ . For each $u \in S$, let $\delta(u)$ denote the set of links covered by u . It is straightforward that

$$\delta(\Gamma) \text{ is the disjoint union of all } E_u \equiv \delta(u) \cap \delta(\Gamma), u \in S. \quad (4.7)$$

We also know from (4.5) and (4.6) that

$$\lim_{n \rightarrow \infty} \Pr\left(|\delta(\Gamma)| = \sum_{v \in \Gamma} d(v) \geq (1 - \epsilon/2) \frac{2n\beta}{\beta + 2}\right) = 1 \quad (4.8)$$

For any node $u \in S$ with nonempty $E_u \cap C_2$, considering any $uv \in E_u \cap C_2$, we see that u and v ($\in \Gamma$) purchase under the same price $p \leq d(v) = \beta$. Since $d(u) \geq \beta + |E_u| \geq p + |E_u \cap C_1| + |E_u \cap C_2|$ and $|\delta(u) \setminus C_2| \geq d(u) - p$, we have $|\delta(u) \setminus C_2 \setminus (E_u \cap C_1)| = |\delta(u) \setminus C_2| - |E_u \cap C_1| \geq |E_u \cap C_2|$. It follows that

For each $u \in S$, there is a subset F_u of $\delta(u) \setminus C_2 \setminus (E_u \cap C_1)$ with $|F_u| = |E_u \cap C_2|$ links.

As F_u is disjoint from both C_2 and $E_u \cap C_1$, we have $c(e) \leq 1$ for any $e \in F_u$, and $c(e) = 0$ for any $e \in F_u \cap E_u$. This enables us to modify c to be another charge function $c' : E \rightarrow \{0, 1, 2\}$ such that $c'(E) = c(E)$ and $c'(e) \leq 1$ for every $e \in \delta(\Gamma)$ as follows. For each $u \in S$, we increase the charge of each link in F_u by 1, and decrease the charge of each link in $E_u \cap C_2$ by 1. The resulting charge c' is as desired because, as (4.7) implies, $\delta(\Gamma) \cap C_2$ is the disjoint union of $E_u \cap C_2$ for all $u \in S$. Therefore, we obtain

$$\text{OPT}(G) = \sum_{p \in \mathbf{p}} p|B(p)| = c(E) = c'(E) \leq 2|E| - |\delta(\Gamma)|.$$

Using (4.8), we have with probability tending to 1 (as $n \rightarrow \infty$)

$$\text{OPT}(G) \leq 2(1 + \epsilon/2)\beta n - (1 - \epsilon/2)\frac{2n\beta}{\beta + 2} \leq \left(2 - \frac{2}{\beta + 2} + \epsilon\right)n\beta.$$

Observing finally that the single pricing with price β obtains revenue $n\beta$ completes the proof. \square

In the special case of $\beta = 1$, Barabási-Albert network G is a tree. The approximation ratio $2 - 2/(\beta + 2) = 4/3$ stands in contrast to the ratio 1.5 in Theorem 4.1 and Remark 4.2.

4.4 Upper and lower bounds for single pricing

Having seen the above constant approximations that single pricing achieves, one may ask: can best single pricing always provide good approximations to optimal iterative pricing for unweighted networks with uniform intrinsic values? The following example shows that, in the worst case, the best single price can only guarantee at most a fraction $1/(\ln \ln n)$ of the optimal revenue.

Example 4.5. *The network G with $n = k(k!) + 1$ nodes consists of $\sum_{i=1}^k i = k(k+1)/2$ node-disjoint cliques and one special node which is adjacent to all other nodes, where the number of $(k!/i)$ -cliques is i for each $1 \leq i \leq k$.*

In the above instance G , there are one node with degree $k(k!)$, which is the special node, and $k!$ nodes with degree $(k!)/i$ for $i = 1, 2, \dots, k$. Recall that $R(p)$ denote the revenue under single pricing (p). Note that $R(k(k!)) = k(k!)$, and $R((k!)/i) = (i(k!) + 1) \cdot (k!)/i = (k!)^2 + (k!)/i$ for $i = 1, \dots, k$. Then the best single price is $k!$, which brings a revenue

$$R(k!) = (k!)^2 + k! = \max_{p \geq 0} R(p).$$

On the other hand the pricing $\mathbf{p} = (p_1, \dots, p_{k+1})$ with $p_1 = k(k!)$, $p_{i+1} = (k!)/i$, $i = 1, \dots, k$, brings revenue $R(\mathbf{p}) = k(k!) + \sum_{i=1}^k (k!)(k!/i - 1) = (k!)^2 \cdot \sum_{i=1}^k (1/i)$. When k tends to infinity,

$$\frac{R(\mathbf{p})}{R(k!)} = \frac{\sum_{i=1}^k \frac{1}{i}}{1 + o(1)} \approx 1 + \ln k = \Theta(\ln \ln n).$$

In complementary to the above example, we show in the following theorem that, with single pricing, one can always assure at least a factor $1/(1 + \ln n)$ of the optimal revenue in unweighed network G with uniform intrinsic values. Let d_1, d_2, \dots, d_n with $d_1 \geq d_2 \geq \dots \geq d_n$ be the degree sequence of G .

Theorem 4.6. $\text{OPT}(G)/\max_{i=1}^n \{id_i\} \leq 1 + \ln n$.

Proof. Since $\sum_{i=1}^n d_i \geq \text{OPT}(G)$, it suffices to show that

$$\max_{i=1, \dots, n} \{id_i\} \geq \sum_{i=1}^n \frac{d_i}{1 + \ln n}.$$

Suppose on the contrary that $id_i < \frac{\sum_{j=1}^n d_j}{1+\ln n}$ for each $1 \leq i \leq n$. Then we have

$$\sum_{i=1}^n d_i < \left(\sum_{i=1}^n \frac{1}{i} \right) \cdot \frac{\sum_{i=1}^n d_i}{1+\ln n} \implies 1+\ln n < \sum_{i=1}^n \frac{1}{i},$$

which is a contradiction. □

5 Conclusion

Our work is an addition to the very limited literature on both pricing with negative network externalities and iterative pricing. The model captures many interesting settings in real-world marketing, and is usually much more challenging than the positive externality counterpart. The hardness result identifies complexity status of a fundamental pricing problem. The algorithms achieve satisfactory performances in general and several important special settings. An interesting direction for future research is to narrow the lower and upper bounds on the approximability of the iterative pricing problem with negative externality. Obtaining more accurate estimations for the optimal pricing is a key to reduce the approximation ratios.

References

- [1] Hessameddin Akhlaghpour, Mohammad Ghodsi, Nima Haghpahan, Vahab S. Mirrokni, Hamid Mahini, and Afshin Nikzad. Optimal iterative pricing over social networks (extended abstract). In *WINE*, pages 415–423, 2010.
- [2] Noga Alon, Yishay Mansour, and Moshe Tennenholtz. Differential pricing with inequity aversion in social networks. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 9–24. ACM, 2013.
- [3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [4] MohammadHossein Bateni, Nima Haghpahan, Balasubramanian Sivan, and Morteza Zadimoghaddam. Revenue maximization with nonexcludable goods. In *Web and Internet Economics*, pages 40–53. Springer, 2013.
- [5] Sayan Bhattacharya, Janardhan Kulkarni, Kamesh Munagala, and Xiaoming Xu. On allocations with negative externalities. In *Proceedings of the 7th International Conference on Internet and Network Economics*, WINE’11, pages 25–36, Berlin, Heidelberg, 2011. Springer-Verlag.
- [6] Francis Bloch and Nicolas Quérou. Pricing in social networks. *Games and Economic Behavior*, 80:243–261, 2013.
- [7] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. The degree sequence of a scale-free random graph process. *Random Structures & Algorithms*, 18(3):279–290, 2001.
- [8] Yann Bramoullé. Anti-coordination and social interactions. *Games and Economic Behavior*, 58(1):30–49, 2007.
- [9] Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135(1):478–494, 2007.
- [10] Ozan Candogan, Kostas Bimpikis, and Asuman Ozdaglar. Optimal pricing in networks with externalities. *Operations Research*, 60(4):883–905, 2012.
- [11] Ning Chen. On the approximability of influence in social networks. *SIAM Journal on Discrete Mathematics*, 23(3):1400–1415, 2009.

- [12] Wei Chen, Pinyan Lu, Xiaorui Sun, Bo Tang, Yajun Wang, and Zeyuan Allen Zhu. Optimal pricing in social networks with incomplete information. In *Internet and Network Economics*, pages 49–60. Springer, 2011.
- [13] Changrong Deng and Saša Pekeč. Money for nothing: exploiting negative externalities. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 361–370. ACM, 2011.
- [14] Joseph Farrell and Garth Saloner. Standardization, compatibility, and innovation. *The RAND Journal of Economics*, pages 70–83, 1985.
- [15] Michal Feldman, David Kempe, Brendan Lucier, and Renato Paes Leme. Pricing public goods for private sale. In *Proceedings of the fourteenth ACM conference on Electronic commerce*, pages 417–434. ACM, 2013.
- [16] Nima Haghpanah, Nicole Immorlica, Vahab Mirrokni, and Kamesh Munagala. Optimal auctions with positive network externalities. *ACM Transactions on Economics and Computation*, 1(2):13, 2013.
- [17] Jason Hartline, Vahab Mirrokni, and Mukund Sundararajan. Optimal marketing strategies over social networks. In *Proceedings of the 17th international conference on World Wide Web*, pages 189–198. ACM, 2008.
- [18] Philippe Jehiel, Benny Moldovanu, and Ennio Stacchetti. How (not) to sell nuclear weapons. *The American Economic Review*, pages 814–829, 1996.
- [19] Michael Katz and Carl Shapiro. Network externalities, competition, and compatibility. *The American economic review*, pages 424–440, 1985.
- [20] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.
- [21] Elchanan Mossel and Sebastien Roch. On the submodularity of influence in social networks. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 128–134. ACM, 2007.
- [22] Roy Radner, Ami Radunskeya, and Arun Sundararajan. Dynamic pricing of network goods with boundedly rational consumers. *Proceedings of the National Academy of Sciences*, 111(1):99–104, 2014.

Appendix: Proof of Theorem 3.7

By reduction from the 3SAT problem, we prove that finding an optimal pricing sequence for the PNC model is NP-hard, even when the underlying network is an unweighted tree without intrinsic values.

A Construction

Let I be an arbitrary instance of the 3SAT problem, whose input is given by n boolean variables x_1, x_2, \dots, x_n , and m clauses $c^j = (x^{j1} \vee x^{j2} \vee x^{j3})$, $1 \leq j \leq m$, where $x^{j\ell}$ is a literal taken from $\{x_1, x_2, \dots, x_n, \neg x_1, \neg x_2, \dots, \neg x_n\}$, $1 \leq j \leq m, 1 \leq \ell \leq 3$. To avoid triviality, we assume that

$$m \geq 5, \text{ and for each } i \in [n], \text{ there exist } j, j' \in [m] \text{ such that } x_i \in c^j \text{ and } \neg x_i \in c^{j'}. \quad (\text{A.1})$$

From the 3SAT instance I , we construct an instance of the PNC problem on tree $G = (V, E)$ with unit weight $w|_E = \mathbf{1}$ and all intrinsic values zero as follows. Let R (resp. \bar{R}) denote the set of ordered pairs (i, j)

such that $i \in [n], j \in [m]$ and $x_i \in c^j$ (resp. $\neg x_i \in c^j$). Clearly, $|R| + |\bar{R}| = 3m$. Let $k = k_0, k_1, k_2, \dots, k_n$ be integers satisfying

$$k_i \geq k_{i-1} + 6m, \text{ for } i = 1, 2, \dots, n, \text{ and } k \geq 6m. \quad (\text{A.2})$$

Tree $G = (V, E)$ has $|V| < 3m(k_n^3) + m(k^2 + k + 1) + (9mk_n^3)$ nodes in total, where V is the disjoint union of the node sets of $3m$ variable gadgets, m clause gadgets and one connection gadget.

- For every $(i, j) \in R$, i.e., $x_i \in c^j$, there is a *variable gadget* $X_i^j = (V_i^j, E_i^j)$ which is a tree rooted at node x_i^j (see Figure 2(a)). Node set V_i^j with $|V_i^j| = k_i^3 - 2k_i^2 + 1 < k_n^3$ is the disjoint union of four sets $\{x_i^j\}, V_{i1}^j, V_{i2}^j$ and V_{i3}^j , where $V_{ih}^j, h = 1, 2, 3$, consists of nodes in X_i^j at distance h from x_i^j .
 - *Literal node* x_i^j , which simulates literal x_i , has degree $k_i - 2$ in X_i^j .
 - V_{i1}^j consists of the $k_i - 2$ neighbors of x_i^j in X_i^j , all having degree k_i .
 - V_{i2}^j consists of $(k_i - 2)(k_i - 1)$ nodes, all having degree $k_i + 1$.
 - V_{i3}^j consists of the $(k_i - 2)(k_i - 1)k_i$ leaves of X_i^j .
- For every $(i, j) \in \bar{R}$, i.e., $\neg x_i \in c^j$, there is a *variable gadget* $\bar{X}_i^j = (\bar{V}_i^j, \bar{E}_i^j)$ which is a tree rooted at node \bar{x}_i^j (see Figure 2(b)). Node set \bar{V}_i^j with $|\bar{V}_i^j| = k_i^2 < k_n^3$ is the disjoint union of three sets $\{\bar{x}_i^j\}, \bar{V}_{i1}^j$ and \bar{V}_{i2}^j , where $\bar{V}_{ih}^j, h = 1, 2$, consists of nodes in \bar{X}_i^j at distance h from \bar{x}_i^j .
 - *Literal node* \bar{x}_i^j , which simulates literal $\neg x_i$, has degree $k_i - 1$ in \bar{X}_i^j .
 - \bar{V}_{i1}^j consists of the $k_i - 1$ neighbors of \bar{x}_i^j in \bar{X}_i^j , all having degree $k_i + 1$.
 - \bar{V}_{i2}^j consists of the $(k_i - 1)k_i$ leaves of \bar{X}_i^j .
- For each clause c^j , there is a *clause gadget* $C^j = (V^j, E^j)$ which is a tree rooted at node c^j (see Figure 2(c)). Node set V^j with $|V^j| = k^2 + k + 1$ is the disjoint union of three sets $\{c^j\}, V^{j1}$ and V^{j2} , where $V^{jh}, h = 1, 2$, consists of nodes in C^j at distance h from c^j .
 - *Clause node* c^j , which simulates the clause, has degree k in C^j .
 - V^{j1} consists of the k neighbors of c^j in C^j , all having degree $k + 1$.
 - V^{j2} consists of the k^2 leaves of C^j .
- For any literal x_i and clause c^j with $x_i \in c^j$, there is a link joining literal node x_i^j and clause node c^j . For any literal $\neg x_i$ and clause c^j with $\neg x_i \in c^j$, there is a link joining literal node \bar{x}_i^j and clause node c^j .
- There is a *connection gadget* $S = (V_S, E_S)$ which is a star centered at node s . The *connection node* s has degree $9mk_n^3 - 1$ in S , and is adjacent to every clause node of G .

Obviously, the above construction of $G = (V, E)$ can be done in polynomial time. It is easy to check that G is a tree. In particular, all the $4m + 1$ node-disjoint gadgets (recall that $|R| + |\bar{R}| = 3m$) are connected by $4m$ links adjacent to clause nodes, where each clause node c^j has exactly four neighbors outside C^j , three being literal nodes and one being the connection node s . Let E_{cl} denote the set of $3m$ links connecting clause nodes and literal nodes, and E_{cs} denote the set of m links connecting clause nodes and connection node s . Then

$$V \text{ is the disjoint union of } \bigcup_{(i,j) \in R} V_i^j, \bigcup_{(i,j) \in \bar{R}} \bar{V}_i^j, \bigcup_{j \in [m]} V^j \text{ and } V_S. \quad (\text{A.3})$$

$$E \text{ is the disjoint union of } \bigcup_{(i,j) \in R} E_i^j, \bigcup_{(i,j) \in \bar{R}} \bar{E}_i^j, \bigcup_{j \in [m]} E^j, E_S, E_{cl}, \text{ and } E_{cs}. \quad (\text{A.4})$$

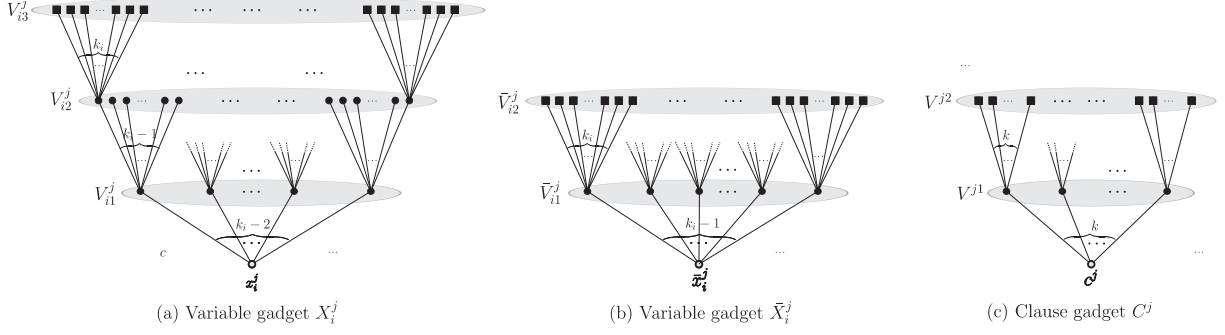


Figure 2: Literal nodes and clause nodes are represented by circles, while auxiliary nodes are represented by solid disks and (pendant) squares.

For any node $v \in V$, let $d(v)$ denote the degree of v in G . The connection node s has degree

$$d(s) = |E_S| + |E_{cs}| = 9mk_n^3 - 1 + m > 9mk_n^3. \quad (\text{A.5})$$

Note that each literal node has exactly one neighbor outside the variable gadget containing it, which is a clause node. Therefore, for any $(i, j) \in R$, we have

$$d(x_i^j) = k_i - 1, \quad d(v) = k_i \text{ for all } v \in V_{i1}^j, \quad \text{and } d(v) = k_i + 1 \text{ for all } v \in V_{i2}^j. \quad (\text{A.6})$$

For any $(i, j) \in \bar{R}$, we have

$$d(\bar{x}_i^j) = k_i, \quad \text{and } d(v) = k_i + 1 \text{ for all } v \in \bar{V}_{i1}^j. \quad (\text{A.7})$$

For every $j \in [m]$, we have

$$d(c^j) = k + 4, \quad \text{and } d(v) = k + 1 \text{ for all } v \in V^{j1}. \quad (\text{A.8})$$

Other nodes, i.e., those not mentioned in (A.5) – (A.8), are exactly leaves of G . It is worthwhile noting from (A.2) that all the non-leaf consumers in the variable gadgets have much larger degrees than the non-leaf consumers in the clause gadgets. This permits us to consider the former consumers before the latter ones.

B Pricing

Given any integer pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ for G , let $\xi = \min\{t : p_t \in \mathbf{p} \text{ and } p_t \leq k + 3\}$ be the first time that the price is equal to or lower than $k + 3$. Note that if $p_1 = d(s)$, then at time 1, only s purchases, and before time ξ , no consumer in the clause gadgets has purchased, i.e.

$$B_{\xi-1}(\mathbf{p}) \cap (\cup_{j=1}^m V^j) = \emptyset. \quad (\text{B.1})$$

Claim 4. Let $D^j = \{x_i^j : x_i \in c^j, i \in [n]\} \cup \{\bar{x}_i^j : \neg x_i \in c^j, i \in [n]\}$. If $p_1 = d(s)$, then the following holds for each $j \in [m]$.

- (i) Either $R_{C^j}(\mathbf{p}) \leq |E^j| + 3$, or $|E^j| + k + 1 \leq R_{C^j}(\mathbf{p}) \leq |E^j| + k + 3$.
- (ii) If $R_{C^j}(\mathbf{p}) \geq |E^j| + k + 1$ then $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap [k + 1, k + 3] \neq \emptyset$.
- (iii) If $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap [k, k + 3] = \{k + 1\}$, then $R_{C^j}(\mathbf{p}) = |E^j| + k + 1$.

Proof. Recall that $\nu_t(v, \mathbf{p})$ denotes the (total) value of the product at time t for consumer $v \in V$. As $p_1 = d(s) = \nu_1(s, \mathbf{p}) > d(v) = \nu_1(v, \mathbf{p})$ for all $v \in V \setminus \{s\}$, only consumer s purchases at time 1. Then for any $t \in [2, \tau]$ it holds that $\nu_t(c^j, \mathbf{p}) \leq \nu_2(c^j, \mathbf{p}) = k + 3$. Thus c^j can only purchase at a price no greater than $k + 3$.

In case of c^j purchasing at some price in $[k + 2, k + 3]$, either all consumers in V^{j1} purchase at some price within $[2, k]$, in which case $R_{C^j \setminus \{c^j\}}(\mathbf{p}) \leq k^2 = |E^j| - k$; or all consumers in $C^j \setminus \{c^j\}$ purchase at price 1, in which case $R_{C^j \setminus \{c^j\}}(\mathbf{p}) = k + k^2 = |E^j|$. So either $R_{C^j}(\mathbf{p}) \leq |E^j| + 3$ or $R_{C^j}(\mathbf{p}) \in \{|E^j| + k + 2, |E^j| + k + 3\}$.

In case of c^j purchasing at price $k + 1$, all consumers in V^{j1} purchase at price $k + 1$, giving $R_{C^j}(\mathbf{p}) = (k + 1)(1 + k) = k^2 + 2k + 1 = |E^j| + k + 1$.

Consider now the case of c^j not purchasing at any price above k . Note first it is possible that c^j never purchases at all under \mathbf{p} , and that consumers in V^{j1} will never purchase at a price higher than $k + 1$. If one and thus all consumers in V^{j1} purchase before c^j , then $R_{C^j}(\mathbf{p})$ is maximized when all consumers in V^{j1} purchase at price $k + 1$, saying $R_{C^j}(\mathbf{p}) \leq (k + 1)k + 3 = |E^j| + 3$. If none of the consumers in V^{j1} purchases before c^j , then $R_{C^j}(\mathbf{p}) \leq \max\{k(1 + k), 1 + k + k^2\} = |E^j| + 1$.

Hence we see that (i) holds, and $R_{C^j}(\mathbf{p}) \geq |E^j| + k + 1$ only if c^j purchases under some price $p_t \in [k + 1, k + 3]$ at time t . Recalling the time point ξ defined at the beginning of this section, we have $t \geq \xi$, which implies $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$. So (ii) is valid.

Suppose now that $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$. Recall from (B.1) that c^j does not purchase before time ξ . It follows that $\nu_\xi(c^j, \mathbf{p}) \geq k + 1$. If $\mathbf{p} \cap [k, k + 3] = \{k + 1\}$, then $p_\xi = k + 1$. It follows that c^j and all consumers in V^{j1} purchase at price $p_\xi = k + 1$, yielding (iii). \square

Note from (A.6) and (A.7) that for any $i \in [n]$, both $\max\{d(v) : (i, j) \in R, v \in V_i^j\}$ and $\max\{d(v) : (i, j) \in \bar{R}, v \in \bar{V}_i^j\}$ are upper bounded by $k_i + 1$; thus none of consumers in $\cup_{j:(i,j) \in R} V_i^j$ and $\cup_{j:(i,j) \in \bar{R}} \bar{V}_i^j$ purchases when the price is above $k_i + 1$. Furthermore, the following two claims can be easily checked by charging (a part of) revenue obtained at a vertex to a subset of edges incident with it, where each edge receives a charge of 1.

Claim 5. For any $(i, j) \in R$, the following hold:

- (i) if $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1, k_i - 1\}$, then $R_{\xi-1, V_i^j}(\mathbf{p}) = |E_i^j| + 1$ and $x_i^j \in B_{\xi-1}(\mathbf{p})$;
- (ii) if $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i\}$, then $R_{\xi-1, V_i^j}(\mathbf{p}) = |E_i^j|$ and $x_i^j \notin B_{\xi-1}(\mathbf{p})$.
- (iii) if $\mathbf{p} \cap [k_i - 1, k_i + 1] \in \{\{k_i + 1, k_i - 1\}, \{k_i\}\}$, then $R_{V_i^j}(\mathbf{p}) \leq |E_i^j| + 1$.

Proof. To see (i), we consider the time, say t , when price $p_t = k_i + 1$ is announced, all consumers in V_{i2}^j purchase, and others in $V_i^j \setminus V_{i2}^j$ do not. We charge the revenue $k_i + 1$ obtained at each consumer of V_{i2}^j to the $k_i + 1$ edges incident with it. Next, at time $t + 1$, price $p_{t+1} = k_i - 1$ is announced, and only x_i^j purchases, because the product value is $k_i - 1$ for x_i^j , and 1 (resp. 0) for each consumer in V_{i1}^j (resp. V_{i3}^j) at that time. Now we charge the $k_i - 2$ edges in E_i^j that are incident with x_i^j . So all edges in E_i^j are charged and 1 revenue is left (this amount corresponds to the edge that connects x_i and the clause gadget C^j), which gives (i), as after x_i^j 's purchase the product values 0 for all consumers in $V_{i1}^j \cup V_{i3}^j$.

To see (ii), note first that only consumers in $V_{i1}^j \cup V_{i2}^j$ purchase under price k_i . For each $v \in V_{i1}^j$, we charge the k_i edges incident with v ; for each $v \in V_{i2}^j$, we charge the k_i pendant edges incident with v . All edges of E_i^j have been charged and no revenue is left. Now, the product values 1 for x_i^j and 0 for all consumers in V_{i3}^j . Hence (ii) holds.

Statement (iii) is straightforward from the proofs of (i) and (ii). \square

Claim 6. For any $(i, j) \in \bar{R}$, the following hold:

- (i) if $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1, k_i - 1\}$, then $R_{\xi-1, \bar{V}_i^j}(\mathbf{p}) = |\bar{E}_i^j|$ and $\bar{x}_i^j \notin B_{\xi-1}(\mathbf{p})$;
- (ii) if $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i\}$, then $R_{\xi-1, \bar{V}_i^j}(\mathbf{p}) = |\bar{E}_i^j| + 1$ and $\bar{x}_i^j \in B_{\xi-1}(\mathbf{p})$.

(iii) if $\mathbf{p} \cap [k_i - 1, k_i + 1] \in \{\{k_i + 1, k_i - 1\}, \{k_i\}\}$, then $R_{\bar{V}_i^j}(\mathbf{p}) \leq |\bar{E}_i^j| + 1$.

Proof. In proving (i), for each consumer in \bar{V}_{i1}^j , we charge the $k_i + 1$ edges incident with it. In proving (ii), for each consumer in \bar{V}_{i1}^j , we charge the k_i pendant edges incident with it; for \bar{x}_i^j , we charge the $k_i - 1$ edges in \bar{E}_i^j that are incident with it. Statement (iii) is then instant. \square

In the rest of this section we discuss the properties of normal pricing sequences.

Claim 7. For any $i \in [n]$, if \mathbf{p} is normal and $\mathbf{p} \cap [k_i - 1, k_i + 1] \notin \{\{k_i + 1, k_i - 1\}, \{k_i\}\}$, then one of the following holds:

- (i) $\mathbf{p} \cap [k_i, k_i + 1] = \emptyset$, in which case $R_{V_i^j}(\mathbf{p}) < |E_i^j| - 12m + 1$ for every every $j \in [m]$ with $(i, j) \in R$, and $R_{\bar{V}_i^j}(\mathbf{p}) < |\bar{E}_i^j| - 12m + 1$ for every every $j \in [m]$ with $(i, j) \in \bar{R}$.
- (ii) $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1\}$ and $\mathbf{p} \cap [2, k_i - 2] \neq \emptyset$, in which case $R_{V_i^j}(\mathbf{p}) \leq |E_i^j| - 6m + 5$ for every $j \in [m]$ with $(i, j) \in R$, and $R_{\bar{V}_i^j}(\mathbf{p}) \leq |\bar{E}_i^j|$ for every $j \in [m]$ with $(i, j) \in \bar{R}$.
- (iii) $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1\}$ and $\mathbf{p} \cap [1, k_i - 2] \subseteq \{1\}$, in which case $R_{V_i^j}(\mathbf{p}) \leq |E_i^j| + 1$ for every every $j \in [m]$ with $(i, j) \in R$, and $R_{\bar{V}_i^j}(\mathbf{p}) \leq |\bar{E}_i^j| + 1$ for every $j \in [m]$ with $(i, j) \in \bar{R}$.

Proof. If $\mathbf{p} \cap [k_i, k_i + 1] = \emptyset$, then no consumer in $(\cup_{j:(i,j) \in R} V_i^j) \cup (\cup_{j:(i,j) \in \bar{R}} \bar{V}_i^j)$ purchases at a price higher than $k_i - 1$. It follows that $R_{V_i^j}(\mathbf{p}) \leq (k_i - 1)|V_{i2}^j \cup V_{i1}^j \cup \{x_i^j\}| = 1 + |E_i^j| - |V_{i2}^j| = |E_i^j| - (k_i - 2)(k_i - 1) + 1$ for every $j \in [m]$ with $(i, j) \in R$ and $R_{\bar{V}_i^j}(\mathbf{p}) \leq (k_i - 1)|\bar{V}_{i1}^j \cup \{\bar{x}_i^j\}| = |\bar{E}_i^j| - |\bar{V}_{i1}^j| = |\bar{E}_i^j| - (k_i - 1)$ for every $j \in [m]$ with $(i, j) \in \bar{R}$. Now $(k_i - 2)(k_i - 1) - 1 > k_i - 1 \geq 12m - 1$, which is implied by (A.2), gives (i).

It remains to consider the case where there exist p_t ($1 \leq t \leq \tau$) that is the maximum price in $\mathbf{p} \cap [k_i, k_i + 1] \neq \emptyset$. If $p_t = k_i$, then the maximality of p_t together with (A.2) and (A.6) – (A.8) implies that all consumers in $V_{i1}^j \cup V_{i2}^j$ with $(i, j) \in R$ and those in $\{\bar{x}_i^j\} \cup \bar{V}_{i1}^j$ with $(i, j) \in \bar{R}$ would purchase under price k_i at time t . Note from (A.5) that consumer s must have purchased by time t . After time t any consumer without the product has value at most $k_{i-1} + 3 < k_i - 1$ (recall (A.6) – (A.8) and (A.2)). It follows from normality of \mathbf{p} that $k_i - 1 \notin \mathbf{p}$, enforcing $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i\}$, a contradiction to the condition $\mathbf{p} \cap [k_i - 1, k_i + 1] \notin \{\{k_i + 1, k_i - 1\}, \{k_i\}\}$ of the claim. Thus $p_t = k_i + 1$, and all consumers in V_{i2}^j with $(i, j) \in R$ and those in \bar{V}_{i1}^j with $(i, j) \in \bar{R}$ purchase under price $k_i + 1$ at time t , bringing about revenues $|E_i^j| - (k_i - 2)$ and $|\bar{E}_i^j|$, respectively. Notice again that s has purchased by time t . After time t , any consumer without the product has value at most $k_i - 1$, which along with the normality of \mathbf{p} gives $k_i \notin \mathbf{p}$. In turn $\mathbf{p} \cap [k_i - 1, k_i + 1] \neq \{k_i + 1, k_i - 1\}$ implies $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1\}$.

As $k_i - 1 \notin \mathbf{p}$, the normality of \mathbf{p} enforces $\mathbf{p} \cap [k_{i-1} + 4, k_i - 1] = \emptyset$. It follows from (A.2) that

$$R_{V_i^j}(\mathbf{p}) \leq |E_i^j| - (k_i - 2) + (k_{i-1} + 3) \leq |E_i^j| - 6m + 5.$$

In case of $\mathbf{p} \cap [2, k_i - 2] \neq \emptyset$, before variable nodes \bar{x}_i^j with $(i, j) \in \bar{R}$ purchase (possibly) at price 1 or 0, all clause nodes have purchased under some price in $\mathbf{p} \cap [2, k_i - 2]$. It follows that all these \bar{x}_i^j with $(i, j) \in \bar{R}$ can only purchase at price 0, yielding (ii).

In case of $\mathbf{p} \cap [2, k_i - 2] = \emptyset$, we have $\mathbf{p} \cap [1, k_i - 2] \subseteq \{1\}$, implying $R_{\bar{V}_i^j}(\mathbf{p}) \leq |\bar{E}_i^j| + 1$ and hence (iii).

Due to the above analysis, it can also be observed that the three situations stated in this claim are all the possible ones. \square

Combining Claims 5(iii), 6(iii) and 7 we obtain the following corollary.

Claim 8. If \mathbf{p} is normal, then $R_{V_i^j}(\mathbf{p}) \leq |E_i^j| + 1$ for all $(i, j) \in R$ and $R_{\bar{V}_i^j}(\mathbf{p}) \leq |\bar{E}_i^j| + 1$ for all $(i, j) \in \bar{R}$.

Claim 9. If \mathbf{p} is normal and $R(\mathbf{p}) > |V|$, then $p_1 = d(s)$.

Proof. Suppose to the contrary that $p_1 \neq d(s)$. By normality of \mathbf{p} , we have $p_1 < d(s)$, and furthermore $p_1 \leq k_n + 1$ (recalling (A.2) and (A.5)–(A.8)). It follows that either $p_1 = 1$, giving $R(\mathbf{p}) = |V|$, or $p_1 \geq 2$, giving

$$R(\mathbf{p}) = R_{\{s\}}(\mathbf{p}) + R_{V \setminus \{s\}}(\mathbf{p}) \leq (k_n + 1) + 2(|E| - |E_S|) = (k_n + 1) + 2(|V| - 1 - |E_S|) < 9mk_n^3 < d(s) < |V|,$$

where the third last inequality uses the fact that $|V| < 3m(k_n^3) + m(k^2 + k + 1) + (9mk_n^3)$ and $|E_S| = 9mk_n^3 - 1$. \square

C Final proof

Having finished all necessary preparations, we are ready to establish the close relation between 3SAT instance I and the PNC instance on tree G .

Theorem 3.7. (Restated) *In the PNC model, computing the optimal pricing sequence is NP-hard, even when the underlying network is an unweighted tree and all the intrinsic values are zero.*

Proof. Let $\text{OPT}(G)$ denote the optimal objective value of the PNC instance on tree $G = (V, E)$. Define

$$L = |E| + (k - 2)m$$

To establish the NP-hardness of the pricing problem, it suffices to prove that $\text{OPT}(G) \geq L$ if and only if the 3SAT instance I is satisfiable.

The “if” part. Suppose that I has a satisfactory truth assignment π with λ variables assigned “TRUE” and the remaining $n - \lambda$ variables assigned “FALSE”. Let $\mathbf{p} = (p_1, p_2, \dots, p_{2n-\lambda+1}, p_{2n-\lambda+2})$ be a solution to the PNC instance on G such that

- $p_1 = d(s)$;
- There are one or two prices for each variable gadget depending on whether the variable is assigned “TRUE” or “FALSE” in π : if x_i is assigned “TRUE” then $k_i \in \mathbf{p}$, if x_i is assigned “FALSE”, then $\{k_i + 1, k_i - 1\} \subset \mathbf{p}$;
- There is a common price for the m clause gadgets: $p_{2n-\lambda+2} = k + 1 \in \mathbf{p}$.

According to Claims 5 and 6, we have

$$\sum_{(i,j) \in R} R_{V_i^j}(\mathbf{p}) + \sum_{(i,j) \in \bar{R}} R_{\bar{V}_i^j}(\mathbf{p}) \geq \sum_{(i,j) \in R} |E_i^j| + \sum_{(i,j) \in \bar{R}} |\bar{E}_i^j|.$$

Furthermore, the satisfiability implies that $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$. Therefore, the condition in Claim 4(iii) holds for every $j \in [m]$, giving $R_{C^j}(\mathbf{p}) = |E^j| + k + 1$ for every $j \in [m]$. It follows from (A.3) that the pricing sequence \mathbf{p} assures a revenue

$$\begin{aligned} R(\mathbf{p}) &= R_{V_S}(\mathbf{p}) + \sum_{(i,j) \in R} R_{V_i^j}(\mathbf{p}) + \sum_{(i,j) \in \bar{R}} R_{\bar{V}_i^j}(\mathbf{p}) + \sum_{j \in [m]} R_{C^j}(\mathbf{p}) \\ &\geq d(s) + \sum_{(i,j) \in R} |E_i^j| + \sum_{(i,j) \in \bar{R}} |\bar{E}_i^j| + \sum_{j \in [m]} (|E^j| + 1 + k) \\ &= |E_S| + |E_{cs}| + \sum_{(i,j) \in R} |E_i^j| + \sum_{(i,j) \in \bar{R}} |\bar{E}_i^j| + \sum_{j \in [m]} (|E^j| + 1 + k) \end{aligned}$$

Now from (A.4) we derive $R(\mathbf{p}) \geq |E| - |E_{cl}| + m(1 + k) = |E| - 3m + m(k + 1) = L$, proving the “if” part.

The “only if” part. Suppose now $\text{OPT}(G) \geq L$. Due to Observation 3.5, there exists a normal pricing sequence $\mathbf{p} = (p_1, p_2, \dots, p_\tau)$ whose objective value $R(\mathbf{p})$ is at least L . As $k \geq 5m$, which implies $L > |E| + 1 = |V|$, we derive from Claim 9 that $p_1 = d(s)$, which validates the subsequent application of Claim 4.

If $R_{C^{j_0}}(\mathbf{p}) < |E^{j_0}| + k + 1$ for some $j_0 \in [m]$, it can be deduced from Claim 4(i) that $R_{C^{j_0}}(\mathbf{p}) \leq |E^{j_0}| + 3$. Recalling (A.3), we derive from Claims 8 and 4(i) that

$$\begin{aligned} R(\mathbf{p}) &= R_{V_S}(\mathbf{p}) + \sum_{(i,j) \in R} R_{V_i^j}(\mathbf{p}) + \sum_{(i,j) \in \bar{R}} R_{\bar{V}_i^j}(\mathbf{p}) + \sum_{j \in [m]} R_{C^j}(\mathbf{p}) \\ &\leq |E_S| + |E_{cs}| + \sum_{(i,j) \in R} (|E_i^j| + 1) + \sum_{(i,j) \in \bar{R}} (|\bar{E}_i^j| + 1) + \sum_{j \in [m] \setminus \{j_0\}} (|E^j| + k + 3) + |E^{j_0}| + 3. \end{aligned}$$

Recalling (A.4), we have

$$\begin{aligned} R(\mathbf{p}) &\leq |E| - |E_{cl}| + (|R| + |\bar{R}|) + (m - 1)(k + 3) + 3 \\ &= |E| - 3m + 3m + (k - 2)m + 5m - k \\ &= L - (5m - k). \end{aligned}$$

Then $k > 5m$ implies $R(\mathbf{p}) < L$, a contradiction. Thus for every $j \in [m]$ we have $R_{C^j}(\mathbf{p}) \geq |E^j| + k + 1$, which along with Claim 4(ii) implies $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ and $\mathbf{p} \cap [k + 1, k + 3] \neq \emptyset$.

Suppose that there exists $i_0 \in [n]$ such that $\mathbf{p} \cap [k_{i_0} - 1, k_{i_0} + 1] \notin \{\{k_{i_0} + 1, k_{i_0} - 1\}, \{k_{i_0}\}\}$. Recall from (A.1) that there exists $j_0 \in [m]$ such that $(i_0, j_0) \in R$. Notice from $\mathbf{p} \cap [k + 1, k + 3] \neq \emptyset$ that $\mathbf{p} \cap [2, k_{i_0} - 2] \neq \emptyset$, because $[k + 1, k + 3] \subseteq [2, k_{i_0} - 2]$ as guaranteed by (A.2). If $\mathbf{p} \cap [k_{i_0} - 1, k_{i_0} + 1] = \{k_{i_0} + 1\}$, then Claim 7(ii) implies that $R_{V_{i_0}^{j_0}}(\mathbf{p}) \leq |E_{i_0}^{j_0}| - 6m + 5$ and further that

$$\begin{aligned} R(\mathbf{p}) &= R_{V_S}(\mathbf{p}) + \sum_{(i,j) \in R} R_{V_i^j}(\mathbf{p}) + \sum_{(i,j) \in \bar{R}} R_{\bar{V}_i^j}(\mathbf{p}) + \sum_{j \in [m]} R_{C^j}(\mathbf{p}) \\ &\leq |E_S| + |E_{cs}| + \sum_{(i,j) \in R \setminus \{(i_0, j_0)\}} (|E_i^j| + 1) + (|E_{i_0}^{j_0}| - 6m + 5) + \sum_{(i,j) \in \bar{R}} (|\bar{E}_i^j| + 1) + \sum_{j \in [m]} (|E^j| + k + 3) \\ &= |E| - |E_{cl}| + (|R| + |\bar{R}| - 1) - 6m + 5 + m(k + 3) \\ &= |E| - 3m + (3m - 1) + m(k - 2) + 5 - m \\ &= L + 4 - m. \end{aligned}$$

Then $m \geq 5$ implies a contradiction to $R(\mathbf{p}) \geq L$, which reduces us to the case $\mathbf{p} \cap [k_{i_0}, k_{i_0} + 1] = \emptyset$ and $R_{V_{i_0}^{j_0}}(\mathbf{p}) < |E_{i_0}^{j_0}| - 12m + 1$ as stated in Claim 7(i). Since $|E_{i_0}^{j_0}| - 12m + 1$ is obviously smaller than $|E_{i_0}^{j_0}| - 6m + 5$, we still have $R(\mathbf{p}) < L$. The contradiction shows that no such an $i_0 \in [n]$ exists, and therefore the conditions in Claims 5 and 6 hold. This enables us to construct a truth assignment π as follows: for each $1 \leq i \leq n$, if $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i + 1, k_i - 1\}$, we assign “FALSE” to variable x_i . Otherwise, that is $\mathbf{p} \cap [k_i - 1, k_i + 1] = \{k_i\}$, we assign x_i “TRUE”. As argued above, $D^j \setminus B_{\xi-1}(\mathbf{p}) \neq \emptyset$ for all $j \in [m]$. Therefore π is indeed a satisfactory truth assignment for I . This completes the “only if” part and the whole proof of Theorem 3.7. \square